

INSTITUT FÜR DEUTSCHE SPRACHE

LDV-INFO 4

Informationsschrift der Arbeitsstelle
Linguistische Datenverarbeitung

1984

Redaktion: S. Dickgießer, P. Mückenmüller

Herausgeber: Institut für deutsche Sprache

ISBN 3-922641-23-7

c Institut für deutsche Sprache 1984

LDV-INFO erscheint jährlich;

Preis pro Heft DM 14.--

Bestellungen sind zu richten an:

Institut für deutsche Sprache

Postfach 5409

D-6800 Mannheim 1

Sylvia Dickgießer

LDV-SERVICE

=====

Wenn Sie durch das vorliegende LDV-Info mit der Arbeitsstelle LDV des IdS gerade erst bekannt gemacht werden, möchte ich Sie auf den LDV-Service aufmerksam machen.

Überlesen Sie bitte einige Zeilen dieses Beitrages, wenn Sie uns vielleicht schon seit 1980 oder 1981 kennen; die Arbeitsstelle hat den Service ja bereits im LDV-Info 1 vom September 1981 vorgestellt.

In der Zwischenzeit wurde ein neuer Rechner installiert (eine Beschreibung der neuen Anlage finden Sie im LDV-Info 3 von 1983), die Softwareausstattung ist verbessert (siehe LDV-Info 3) und die Datenbasis des IdS erheblich erweitert worden.

Der erste Abschnitt meines Beitrages enthält einen Überblick über das zur Zeit für den Service verfügbare Datenmaterial. Im zweiten Abschnitt werde ich einen Teil unserer Arbeiten für externe Wissenschaftler dokumentieren.

Die Hinweise auf Projekte externer Wissenschaftler, für die von der Arbeitsstelle LDV Serviceaufträge ausgeführt wurden, sollen u.a. die Vielfalt der LDV-unterstützten Forschungsarbeiten verdeutlichen, die Adressen eine Kontaktaufnahme ermöglichen, wenn Sie Näheres über die Verwendung der Daten erfahren möchten.

1. Die Arbeitsstelle LDV und ihre Daten

Die Arbeitsstelle LDV des IdS besteht seit 1980. Zu den Aufgaben des Teams (Stand 1984: 2,5 wissenschaftliche Mitarbeiter, 3 Programmierer, 1 Systemprogrammierer, 2 Operateure, 1 Datenerfasserin) gehören u.a. die Bereitstellung maschinenlesbarer Textkorpora zur geschriebenen und gesprochenen deutschen Gegenwartssprache, die Korpusaufbereitung z.B. in Form von Registern, die Entwicklung von Textaufbereitungs- und Analyseprogrammen (vgl. Wolfgang Teubert, Die Arbeitsstelle Linguistische Datenverarbeitung, LDV-Info 3, 1983).

Die Arbeitsstelle unterstützt in erster Linie die Forschungsarbeiten innerhalb des IdS, übernimmt jedoch auch solche Serviceaufträge externer Wissenschaftler, die im Rahmen der gegebenen Personalkapazitäten erledigt werden können. Schreiben Sie uns, wenn Sie meinen, wir sollten für Ihr Projekt tätig werden. Wir werden prüfen, in welcher Art und in welchem Umfang wir helfen können.

Leider ist es uns nicht möglich, Serviceaufträge kostenlos auszuführen. Die Kostenberechnung erfolgt nach Maßgabe der bestehenden Preisliste bzw. Gebührenordnung (siehe Anhang), sofern es sich um die Weitergabe vorhandener Daten oder die Anwendung von Standardprogrammen handelt. In der Regel wird nur die für die Auftragserledigung gebrauchte Rechenzeit abgerechnet. Bei speziellen, aufwendigen Programmierarbeiten entstehen jedoch zusätzliche Kosten. Im Falle sehr arbeitsintensiver Aufträge, die wegen hoher Arbeitsauslastung von LDV-Mitarbeitern nicht erledigt werden können, sehen wir zwei mögliche Verfahrensweisen. Wir laden Sie entweder ein, nach Mannheim zu kommen und gegen eine Vergütung der Rechenzeit die Anlage des IdS zu nutzen, oder wir bemühen uns, geeignete Personen zu finden, die Ihre Arbeiten auf Werkvertragsbasis übernehmen. Die Kosten für Werkverträge müssen wir den Auftraggebern leider in vollem Umfang anlasten.

Im Anschluß finden Sie eine Übersicht über die für den Service verfügbaren Text- und Wortdaten des IdS.

Korpora

1. Mannheimer Korpus 1 (MK1), geschriebene deutsche Gegenwartssprache, ca. 2.2 Mio. laufende Wortformen
2. Mannheimer Korpus 2 (MK2), geschriebene deutsche Gegenwartssprache, ca. 0.3 Mio. laufende Wortformen
3. Bonner Zeitungskorpus, ost- und westdeutsche Gegenwartssprache, ca. 3 Mio. laufende Wortformen
4. Freiburger Korpus (FK), gesprochene deutsche Gegenwartssprache, ca. 0.5 Mio. laufende Wortformen
5. Dialogstrukturenkorpus (DSK), gesprochene deutsche Gegenwartssprache, ca. 2 Mio. laufende Wortformen. Die Texte sind z.T. identisch mit denen des FK, liegen jedoch in einer anderen Transkription vor.

Alle Korpora sind in zeilen- und satznumerierter Form verfügbar (DSK: Numerierung der Sprecheräußerungen) und an die Textdatenbasis des Programms REFER angeschlossen.

Mit REFER können Häufigkeiten von Wortformen und Wortformenbestandteilen bzw. Kombinationen von Wortformen und Wortformenbestandteilen ermittelt werden. Das Programm druckt Textbelege aus und ermöglicht die Erstellung von KWIC-Indizes (KWIC = keyword in context). Ausführliche Programmbeschreibungen enthalten die LDV-Infos 1 und 3.

Register

Neben KWIC-Indizes erstellen wir Wortformenregister mit Belegstellenangaben (tokens) und Häufigkeitsregister (types).

Bereits vorhanden ist ein maschinenlesbares VALENZREGISTER mit morphosyntaktischen Angaben zu ca. 4000 ausgewählten Verben des MK1.

Lexika

Eine vorläufige Version des MORPHOLOGISCHEN VOLLFORMENLEXIKONS (MOLEX) umfaßt ca. 1.5 Mio. Einträge mit morphosyntaktischen Angaben zu Nomen, Adjektiven und Verben.

Die maschinenlesbaren Basisdaten der RÜCKLÄUFIGEN WORTLISTE (RWL) umfassen mehr als 153 000 nach Endungen sortierte Grundformen mit Angaben zur Wortart.

Das VALENZLEXIKON enthält morphosyntaktische Angaben zu ca. 400 ausgewählten Verben des Grundwortschatzes.

Die genannten Korpora, Register und Lexika können in maschinenlesbarer Form gekauft oder gemietet werden. Wir geben die Daten natürlich auch in gedruckter Form weiter.

2. Serviceaufträge

Von den Daten und Serviceleistungen der Arbeitsstelle LDV haben in der Zeit von Januar 1982 bis März 1984 mehr als 100 externe Wissenschaftler Gebrauch gemacht.

60 Aufträge betrafen

- a) die Belegstellensuche
- b) die Registererstellung
- c) das Kopieren maschinenlesbarer Texte und Korpora
- d) das Ausdrucken von Texten der Mannheimer Korpora
- e) das Ausdrucken von Optionen des Valenzregisters
oder Valenzlexikons

In einigen Fällen erhielten wir keine Informationen über die Verwendung des gelieferten Materials. 42 Aufträge sind mit Projektangabe auf den folgenden Seiten dokumentiert.

a) Belegstellensuche

***** * Auftraggeber *****	***** Projekt *****	***** Auftrag *****	***** Ausgangsdaten *****
Prof. Dr. W. L-Bald, RWTH Aachen, Institut für Anglistik, Karmannstraße 17-19, 5100 Aachen	Die Verwendung der englischen Form "one" in Vergleich mit den Formen "ein-" und "man" in der gesprochenen Sprache	Ausdruck von Belegstellen für die Wortform "man"	FK
Ulrike Rastert, Sandweg 11, 4790 Paderborn	Dissertation: "Zur Darstellbarkeit von Modalpartikeln im Wörterbuch, am Beispiel von "doch" und "ja"	Ausdruck von Belegstellen für ausgewählte Adverbien	MK1
Friedhelm Beck, c/o Hossenfelder, Argelanderstraße 138, 5300 Bonn 1	Juristische Dissertation, Wortanalyse: "Überwachen" in § 87 I 6 BetrVG	Ausdruck von Belegstellen für "Überwachen"	MK1
Prof. Dr. Broder Carstensen, Universität Paderborn, Fachbereich 3, Warburger Straße 100, 4790 Paderborn	Wörterbuch der Anglizismen im gegenwärtigen Deutsch	Suche nach Belegstellen für ausgewählte Substantive und alle Wortformen mit der Endung "ical"	MK1, MK2, BZK, FK
Prof. Dr. Si Ho Chong, German Department, Teachers College, Kyung-Pook National University, Daegu 635, Republik Korea	Wortfeldanalyse: Verben des Veranlassens	Ausdruck von Belegstellen für ausgewählte Verben	MK1

a) Belegstellersuche

 * Auftraggeber

 ***** Projekt
 ***** Auftrag
 ***** Ausgangsdaten

Prof. Dr. Jürgen
 Dittmann, Albert
 Ludwigs-Universität,
 Deutsches Seminar, Aht-
 für Sprache und Ältere
 Literatur,
 Werthmannplatz, 7800
 Freiburg

Der Gebrauch des Futurs
 in der gesprochenen
 deutschen Standardsprache

Ausdruck von
 Belegstellen für die
 Wortformen "wirst",
 "wird" und "werd-"

FK

Dr. W. Dodd, Department
 of German, University
 of Birmingham, PO Box
 363, Birmingham B31
 2AU, England

Wortfeldanalyse: Verben
 der Körperbewegung

Ausdruck von
 Belegstellen für
 ausgewählte Verben

MK1, FK

Charles J. Fillmore, Dep.
 of Linguistics,
 University of
 California, Berkeley,
 California 94720, USA

"Minor Grammatical
 Constructions"

Ausdruck von
 Belegstellen für die
 Wendung "geschweige
 denn"

MK1

Josef Garncarz,
 Roonstraße 38, 5000
 Köln

Dissertation:
 "Möglichkeiten der
 Selbstaufklärung in der
 ästhetischen Rezeption
 von Naivität"

Ausdruck von
 Belegstellen für die
 Wortformen "naiv-"
 und "Naiv-"

MK1

Jorunn Hetland, More og
 Romsdal
 Distrikthøgskole,
 N-6101 Volda, Norwegen

Dissertation:
 "Satzadverbale im
 modernen Deutsch"

Ausdruck von
 Belegstellen für das
 Suffix "weise"

MK1

a) Belegstellensuche

***** * Auftraggeber *****	***** Projekt *****	***** Auftrag *****	***** Ausgangsdaten *****
Irma Hyvärinen, Yläkartanonkuja 2A, SF-02360 Espoo 36, Finnland	Dissertation: "Der Infinitiv in der Funktion des Prädikateils, des Subjekts und des Objekts im Finnischen und im Deutschen	Ausdruck von Belegstellen für ACI-Konstruktionen	MK1
Werner Kerber, Universität Trier, Fachbereich II, Postfach 3875, 5500 Trier	Wörterbuchartikel: Semantische Analyse des Begriffs Job	Ausdruck von Belegstellen für die Wortformen "Job" und "-job"	MK1, MK2
Dr. Klaus-Michael Köpcke, Universität Hannover, FB Erziehungswissenschaften n 1, Bismarckstraße 2, 3000 Hannover	Probleme der Genuszuweisung bei Nomen der deutschen Gegenwartssprache	Ausdruck von Belegstellen für ausgewählte Nomen	MK1
Karl-Heinz Lukassen, Mittelweg 120, 4190 Kleve	Dissertation: "Verschmelzungsformen von Präpositionen und Artikeln"	Ausdruck von Belegstellen für ausgewählte Kombinationen von Präpositionen und Artikeln	MK1
Prof. Dr. Inger Rosengren, Lunds Universität, Tyska Institutionen, Helgonabacken 14, S-22362 Lund, Schweden	Aufsatz: "Der deutsche Finalsatz in kontrastiver Sicht", in Studia Linguistica 37:2, 1983	Ausdruck von Belegstellen für "Komma" und "damit" auf Karteikarten	MK1, FK

a) Belegstellensuche

 * Auftraggeber ***** Projekt ***** Auftrag ***** Ausgangsdaten *

Dr. Danuta Rytel-Kuc,
 Akademie der
 Wissenschaften,
 Warszawa, Polen

Habilitation:
 "Passivsätze im
 Deutschen, Polnischen und
 Tschechischen"

Ausdruck von
 Belegstellen für
 passivische
 Konstruktionen nach
 Vorgabe von
 Satznummern auf
 Karteikarten

MK1

Prof. Dr. Laurits
 Saltveit, Asvn. 10,
 N-1400 Ski, Norwegen

Variative Grammatik

Ausdruck von
 Belegstellen für
 Nebensatzvarianten

MK1, FK

Dr. Wilfried Seibicke,
 Universität Heidelberg,
 Germanistisches
 Seminar, Postfach
 105760, 6900 Heidelberg

Sprachformen der Ironie.
 Analyse des tatsächlichen
 Gebrauchs der Wörter
 "Ironie" und "ironisch"
 in der deutschen
 Gegenwartssprache

Ausdruck von
 Belegstellen für die
 Wortformen "Ironie"
 und "ironisch"

MK1

Juliane Stegner,
 Westendriederstraße 27,
 8000 München 2

Ellipse als Mittel zum
 Ausdruck der
 Thema-Rhema-Struktur,
 insbesondere Vorwärts-
 und Rückwärtsellipsen in
 mit "und" koordinierten
 Ausdrücken

Ausdruck von
 Belegstellen für die
 Kombinationen "und
 ... hat ... ge-..."
 und "und ... hat ...
 be-..."

FK

Michael Sokolowski,
 Bergstraße 63, 6900
 Heidelberg

Dissertation:
 "Kompetetitives Handeln im
 politischen
 Sprachgebrauch"

Ausdruck von
 Belegstellen für
 ausgewählte Verben
 und Substantive

MK1

a) Belegstellensuche

 * Auftraggeber ***** Projekt ***** Auftrag ***** Ausgangsdaten *****

Dr. Folke Freund, Dr.
 Birger Sundqvist,
 Uppsala Universitet,
 Tyska Institutionen,
 Box 513, S-75120
 Uppsala, Schweden

Schwedisch-deutsche
 Universitätsgrammatik

Ausdruck von
 Belegstellen für die
 Wortform "damit"

MK1,FK

Maurice Vliegen,
 Katholieke
 Universiteit, Instituut
 Nuits Kamer 9.02a,
 Erasmusplein 1, Postbus
 9103, NL-6500 Nijmegen,
 Niederlande

Dissertation:
 "Semantische und
 syntaktische Beschreibung
 der Verben der auditiven
 Wahrnehmung im
 Niederländischen und im
 Deutschen

Ausdruck von
 Belegstellen für
 ausgewählte Verben
 der auditiven
 Wahrnehmung

MK1

Wolfgang Zimmermann,
 Friedrich
 Alexander-Universität,
 Deutsches Seminar,
 Bismarckstraße 1, 8520
 Erlangen

Dissertation:
 "Syntaktisch-semantisch-p
 ragmatische Analyse der
 Verben des Fragens"

Ausdruck von
 Belegstellen für
 ausgewählte Verben
 des Fragens

MK1,FK

Zorica Sergijevska,
 Franklin Ruzvelt, 6/21,
 YU-9100 Skopje,
 Jugoslawien

Partikel mit illokutiver
 Funktion im gesprochenen
 Deutsch und ihre
 Entsprechungen im
 gesprochenen
 Makedonischen

Ausdruck von
 Belegstellen für
 ausgewählte Partikel
 und jeweils 2
 Kontextsätzen auf
 Karteikarten

FK

b) Registererstellung

***** * Auftraggeber *****	***** Projekt *****	***** Auftrag *****	***** Ausgangsdaten *****
Prof. Dr. W. D. Bald, RWTH Aachen, Institut für Anglistik, Karmannstraße 17-19, 5100 Aachen	Die Verwendung der englischen Form "one" im Vergleich mit den Formen "ein-" und "man" in der gesprochenen Sprache	KWIC-Index für die Wortform "ein"	FK
Rolf Fehrens, Kepler Gymnasium, Habsburgerstraße 103, 7800 Freiburg	Unterrichtseinheit: Textverschlüsselung und Textentschlüsselung	Häufigkeitsregister der Buchstaben der deutschen Sprache, statistische Auswertungen des MK1	MK1
Jeanette Huber, Compulearn LTD., Software für den Unterricht, Sudetenstraße 3/Wasserturm, 6080 Gross-Gerau	Sprachlehrprogramme für den Erwachsenenunterricht, Einsatz computergesteuerter Übungseinheiten	Register der 5000 häufigsten Wortformen des FK	FK
Dr. Klaus-Michael Köpcke, Universität Hannover, FB Erziehungswissenschaften n 1, Bismarckstraße 2, 3000 Hannover1	Problematik der Genuszuweisung bei den Nomen der deutschen Gegenwartssprache	KWIC-Index für die Wortformen "Teil" und "-teil", KWIC-Indizes für ausgewählte Verben	MK1, FK
Dr. C. L. Naumann, RWTH Aachen, Ahornstraße 55, 5100 Aachen	Handreichung für Lehrer an Grundschulen: Erläuterung der Konstruktion eines Rechtschreib-Grundwortsch atzes durch die Vorgabe von Wortlisten mit Häufigkeitsangaben	Register der 5000 häufigsten Wortformen des MK1, alphabetisch und nach Häufigkeiten sortiert	MK1

b) Registererstellung

***** * Auftraggeber *****	***** Projekt *****	***** Auftrag *****	***** Ausgangsdaten *****
Nicolai Petersen, Tatvaskalantie 1 A 2, SF-00600 Helsinki 60, Finnland	Die Beziehung zwischen phonologischen, morphologischen und semantischen Merkmalen bei verschiedenen Lexemgruppen im Deutschen	Register aller Wortformen mit einer Frequenz größer/gleich 5 im MK1	MK1
Ursula Ratzke, Arp-Schneider-Stieg 53, 2000 Hamburg 96	Magisterarbeit: "Das Wort "Geist" in deutschsprachigen Romanen um 1800 und in der Gegenwart	KWIC-Indizes für die Wortform "Geist" und Komposita	MK1
Dr. Folke Freund, Dr. Birger Sundqvist, Uppsala Universitet, Tyska Institutionen, Box 513, S-75120 Uppsala, Schweden	Schwedisch-deutsche Universitätsgrammatik	KWIC-Indizes für alle Flexionsformen von "Paar", "Menge", "Dutzend" und "Preis"	MK1, MK2, FK
Martin Wind, van Eysingenlaan 187, NL-3527 Utrecht, Niederlande	Programm zur Umwandlung graphischer in phonetische Zeichen	Register der 100000 häufigsten Wortformen des MK1 auf Magnetband	MK1
Horst Zuse, TU Berlin, FB 20, Informatik, Straße des 17. Juni 135, 1000 Berlin 12	Diplomarbeit: "Qualitative Bewertung von Wortstammalgorithmen"	Präfix-, Suffix- und Flexionssuffixregist er	MOLEX

c) Kopieren maschinenlesbarer Texte und Korpora

 * Auftraggeber
 ***** Projekt
 ***** Auftrag
 ***** Ausgangsdaten

Prof. Dr. Wolfgang
 Herlitz, Muntstraße 4,
 NL-3572 Utrecht,
 Niederlande

Kontextspezifische
 Analysen der
 syntaktisch-semantischen
 Normalformen der
 häufigeren deutschen
 Verben

Kopieren von Texten
 des MK1

MK1

Kari Helene Hestvik,
 Universitet Oslo,
 Germanistik Institutt,
 Postboks 1004 -
 Blindern, Oslo 3,
 Norwegen

Examensarbeit: "Das
 Verhältnis von
 Zustandspassiv und
 Zustandsreflexiv bei
 emotiven Verben unter
 besonderer
 Berücksichtigung etwaiger
 semantisch-pragmatischer
 Interpretationsmöglichkei
 ten"

Kopieren des MK1

MK1

Kenate Kempf, 4471 R
 Hamilton Ave., CA 95130
 San Jose, USA

Dissertation:
 "Traditionelle und
 computerunterstützte
 Methoden der Textanalyse

Kopieren des BZK

BZK

d) Ausdrucken von Texten

 * Auftraggeber ***** Projekt ***** Auftrag ***** Ausgangsdaten *****

Ulrike Paster, Sandweg
 11, 4790 Paderborn

Dissertation: "Zur
 Darstellbarkeit von
 Modalpartikeln im
 Wörterbuch, am Beispiel
 von "doch" und "ja"

Ausdruck von
 Textausschnitten des
 MK1

MK1

Prof. Dr. Jürgen
 Pittmann, Albert
 Ludwigs-Universität,
 Abt. für Sprache und
 Ältere Literatur,
 Werthmannplatz, 7800
 Freiburg

Der Gebrauch des Futurs
 in der gesprochenen
 deutschen Standardsprache

Ausdruck von Texten
 des FK, in denen die
 Wortformen "wirst",
 "wird" und "werd-"
 vorkommen

FK

John Fowler, The College
 Language Laboratory,
 University College of
 Swansea, Singleton
 Park, Swansea SA2 8PP,
 Wales

Konversationsanalytischer
 Vergleich gesprochener
 deutscher und englischer
 Standardsprache, Einsatz
 von Bandaufnahmen
 gesprochener deutscher
 Standardsprache als
 Lehrmittel

Ausdruck von Texten
 des DSK

DSK

Konrad Kärn, Im Winkel 9,
 7800 Freiburg

Pragmatik in der Poetik -
 "Destroyer" als Kategorie
 in Quasitranskripten und
 Transkripten

Ausdruck eines
 Textes des FK.

FK

e) Ausdrucken von Optionen des Valenzregisters

 * Auftraggeber ***** Projekt ***** Auftrag ***** Ausgangsdaten *****

Valenzregister

Verb-Valenz-Lexikon für
 den Sprachgebrauch bei
 Hör-Sprachgeschädigten im
 Vergleich zu Hörenden

Dr. Klaus-B. Günther,
 Pädagogische Hochschule
 Heidelberg,
 Forschungsstelle für
 Angewandte
 Sprachwissenschaft zur
 Rehabilitation
 Behinderter, Im
 Neuenheimer Feld 561,
 6900 Heidelberg